

0540

10/09 OIPE

## RAW SEQUENCE LISTING

DATE: 06/08/2001

PATENT APPLICATION: US/09/840,762

TIME: 14:15:18

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\I840762.raw

**ENTERED**

```

3 <110> APPLICANT: Vreeland, Valerie
4     Ng, Kwan L.
5     The Regents of the University of California
7 <120> TITLE OF INVENTION: Recombinant Vanadium Haloperoxidases and Their Uses
9 <130> FILE REFERENCE: 023070-087100US
11 <140> CURRENT APPLICATION NUMBER: 09/840,762
12 <141> CURRENT FILING DATE: 2001-04-23
14 <150> PRIOR APPLICATION NUMBER: 09/151,189
15 <151> PRIOR FILING DATE: 1998-09-10
17 <160> NUMBER OF SEQ ID NOS: 11
19 <170> SOFTWARE: PatentIn Ver. 2.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2931
23 <212> TYPE: DNA
24 <213> ORGANISM: Fucus distichus
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (228)..(2258)
29 <223> OTHER INFORMATION: vanadium bromoperoxidase
31 <400> SEQUENCE: 1
32 cgcgacaag ccttggaaga gaggttgccc aattcaacag agcgaggccc gtgaaggtgt 60
34 ggaggacacg tgctacaagc tgatccacga gagcctcaac ttccctactg atacgggagt 120
36 ttgtactgcg ccgcggttgcc aaaaaccgca actttaacaa gcgctcgaga gcgccacatg 180
38 cttcccacgc atccacaaaa tcgacagtgg tatcgctgag cttgaat atg ctt tgc 236
39                                     Met Leu Cys
40                                     1
42 cat gca gcg gac acg aca aga ggc tct cct atg cct gac acc gga gtg 284
43 His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp Thr Gly Val
44     5                               10                               15
46 ctt cgg ttg ctc aca tca gag cag cgc gct aaa ggt tgg aga cgc cag 332
47 Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp Arg Arg Gln
48 20                               25                               30                               35
50 tta gag ggg gag aaa tca cta ggt ttt cat cca agc gag acg cct tat 380
51 Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu Thr Pro Tyr
52     40                               45                               50
54 atc aag tac ttg gaa ggc tct gag act tgg aag aag gtt aag ctt cca 428
55 Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val Lys Leu Pro
56     55                               60                               65
58 acg gac ggc ata tcg gct tcc aag atc ctg ggt aaa att atg gcc agg 476
59 Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile Met Ala Arg
60     70                               75                               80
62 gtc cgc atc gct acc gcc ttg gct gtg gta ctg gcc gca ccc tgt ttg 524
63 Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala Pro Cys Leu
64     85                               90                               95
66 gca ttc gac gag gtc aca gcc agt ggt gtt ttc cct gag gaa cac aag 572
67 Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu Glu His Lys
68 100                               105                               110                               115

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,762

DATE: 06/08/2001

TIME: 14:15:18

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\I840762.raw

```

70 cac acc ggg gag gga aga cac ctc cag acc tgt aca aac tcc gac gat 620
71 His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn Ser Asp Asp
72 120 125 130
74 gcg ctg gat ccg acg gcg ccg aat aga agg gac aac gta gct ttt gcg 668
75 Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val Ala Phe Ala
76 135 140 145
78 tcg cgg cgc gat gcc gcc agg cga gaa cgt gac ggg aca ggg act gtc 716
79 Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr Gly Thr Val
80 150 155 160
82 tgc caa atc act aac gga gaa act gat ttg gct acc atg ttc cac aag 764
83 Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met Phe His Lys
84 165 170 175
86 tct ctg cca cac gat gaa ctg gga cag gta acc gca gac gac ttc gct 812
87 Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp Asp Phe Ala
88 180 185 190 195
90 atc ctc gag gac tgc atc tta aac gga gat ttc agc att tgc gag gac 860
91 Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile Cys Glu Asp
92 200 205 210
94 gtg cct gcg gga gac ccg gcg ggt cgc ctc gtc aat cct acc gct gcg 908
95 Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro Thr Ala Ala
96 215 220 225
98 ttt gcc atc gac ata tcc ggt ccc gca ttc tgc gct acg aca ata ccc 956
99 Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr Thr Ile Pro
100 230 235 240
102 ccg gta cct acc ctt tcc tct cct gag ctc gcc gct cag ttg gcg gag 1004
103 Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln Leu Ala Glu
104 245 250 255
106 cta tac tgg atg gcg ctg gcc agg gat gta ccc ttt atg cag tat ggc 1052
107 Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met Gln Tyr Gly
108 260 265 270 275
110 acc gac gaa att acc act acc gcg gca gcc aac ctc gct gga atg gga 1100
111 Thr Asp Glu Ile Thr Thr Thr Ala Ala Ala Asn Leu Ala Gly Met Gly
112 280 285 290
114 ggc ttc cca aat ctg gac gcc gtg tgc ata ggg tcc gat ggt acg gtg 1148
115 Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp Gly Thr Val
116 295 300 305
118 gac ccg ttc tcc cag ctc ttc cga gcg acc ttc gtt ggt gtt gaa acg 1196
119 Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly Val Glu Thr
120 310 315 320
122 ggg ccc ttt gtc tct cag ctg ctc gtg aac agc ttc acc atc gac gct 1244
123 Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr Ile Asp Ala
124 325 330 335
126 att acg gtc gaa ccg aag cag gag aca ttc gcc ccc gac ttg aac tat 1292
127 Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp Leu Asn Tyr
128 340 345 350 355
130 atg gtc gat ttt gac gaa tgg ctg aac att cag aat ggt gga ccc ccg 1340
131 Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly Gly Pro Pro
132 360 365 370
134 gcc ggc ccc gaa gag tta gac gaa gag ctg cgt ttt atc cgt aac gcc 1388

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,762

DATE: 06/08/2001

TIME: 14:15:18

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\I840762.raw

```

135 Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile Arg Asn Ala
136          375          380          385
138 cgc gac ctg gcc agg gtc tcc ttc gtg gac aat atc aac acc gaa gct 1436
139 Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn Thr Glu Ala
140          390          395          400
142 tat cgc ggg tct ctt atc cta ctt gag ctg gga gcc ttc agc agg ccc 1484
143 Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe Ser Arg Pro
144          405          410          415
146 ggt atc aac ggt cca ttc atc gac agt gat cgg cag gcg ggc ttc gtc 1532
147 Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala Gly Phe Val
148 420          425          430          435
150 aac ttc ggc acg tct cac tac ttc aga ttg ata ggt gcc gcc gag ctg 1580
151 Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala Ala Glu Leu
152          440          445          450
154 gcg cag cgt gcc tcg tgt tac caa aag tgg cag gtg cat cga ttt gca 1628
155 Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His Arg Phe Ala
156          455          460          465
158 cgc ccc gag gct ctc ggg ggt acc ctc cac aac acc atc gcg ggg gat 1676
159 Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile Ala Gly Asp
160          470          475          480
162 cta gat gca gac ttc gac atc tcc ctt ctt gaa aat gat gag ctc ttg 1724
163 Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp Glu Leu Leu
164          485          490          495
166 aaa cgt gtg gcg gag ata aat gcg gcg cag aat ccc aac aac gag gtc 1772
167 Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn Asn Glu Val
168 500          505          510          515
170 acc tac ctt ctt cca caa gct atc caa gtg gga tcg cca acg cac cct 1820
171 Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro Thr His Pro
172          520          525          530
174 tcc tac ccg tcc ggc cac gct acc caa aat gga gca ttt gcc aca gtt 1868
175 Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe Ala Thr Val
176          535          540          545
178 ctg aag gcc ctc att ggc cta gat cgg gga ggt gag tgc ttc cct aac 1916
179 Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys Phe Pro Asn
180          550          555          560
182 ccc gtg ttc cca agc gat gac ggc ctg gaa cta atc aac ttc gaa ggg 1964
183 Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn Phe Glu Gly
184          565          570          575
186 gca tgc ctt aca tat gag gga gag atc aac aag ctc gcg gtc aac gtc 2012
187 Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala Val Asn Val
188 580          585          590          595
190 gca ttt ggg agg cag atg ctg ggc atc cac tat cgg ttc gac ggt atc 2060
191 Ala Phe Gly Arg Gln Met Leu Gly Ile His Tyr Arg Phe Asp Gly Ile
192          600          605          610
194 caa ggc cta ctt ctc gga gag aca atc act gta cga aca ctt cac cag 2108
195 Gln Gly Leu Leu Leu Gly Glu Thr Ile Thr Val Arg Thr Leu His Gln
196          615          620          625
198 gag ctg atg acg ttc gcc gag gaa gcc acc ttt gaa ttc cgc tta ttc 2156
199 Glu Leu Met Thr Phe Ala Glu Glu Ala Thr Phe Glu Phe Arg Leu Phe

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,762

DATE: 06/08/2001

TIME: 14:15:18

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\I840762.raw

```

200          630          635          640
202 acc gga gag gtc atc aaa ctt ttc cag gac ggg aca ttc tcc atc gat 2204
203 Thr Gly Glu Val Ile Lys Leu Phe Gln Asp Gly Thr Phe Ser Ile Asp
204          645          650          655
206 gga gat atg tgt tcc ggt ttg gtt tac act ggc gtg gcg gac tgc cag 2252
207 Gly Asp Met Cys Ser Gly Leu Val Tyr Thr Gly Val Ala Asp Cys Gln
208 660          665          670          675
210 gct tagtgcagaa aataataatt gtcggatgct taaaatgcac ccacgaccaa 2305
211 Ala
213 gtcgtcgagt cacgtcgccg gagcatcctt cagcgaaaaa ggagagtaac ctatatgcta 2365
215 tagaggagaa ccacggagta caatgcaggt tcttttacc a tgtacattgg attgcagtaa 2425
217 gtgcgggttag agagggatac gttaaacgtg cttgcctgtg tatatgatac atttgcacatg 2485
219 gaaatattag aatgcgttga cttgacttca ccatgaaata ccatgatcgc gtgggtgtgct 2545
221 gctttcacct gtcggagcgg tacgtaagat gtgctttcta ctgagccgtt tgtgtttagt 2605
223 ccattccgcg tggcagtgta aacaaagagg atgtagtctc gccctcagtt tggagagtac 2665
225 cgtaggtggc aggacgtata tctctggtag cggctctgta agaacttcca caagaccgtt 2725
227 tacgttttgt tgtttagtgc atgcctcttc gttacttgac cgatccattg agagtacctg 2785
229 taccagtatg gtgtaagaca ttttttctc ctggttatgga tctgtagaac agctaggtgt 2845
231 tgttttatac acaggatgct ataaaatagg gatgttgata atggcatcgg tactcatgaa 2905
233 accgcaaaat ggcgatagat attccc 2931
236 <210> SEQ ID NO: 2
237 <211> LENGTH: 676
238 <212> TYPE: PRT
239 <213> ORGANISM: Fucus distichus
241 <400> SEQUENCE: 2
242 Met Leu Cys His Ala Ala Asp Thr Thr Arg Gly Ser Pro Met Pro Asp
243 1 5 10 15
245 Thr Gly Val Leu Arg Leu Leu Thr Ser Glu Gln Arg Ala Lys Gly Trp
246 20 25 30
248 Arg Arg Gln Leu Glu Gly Glu Lys Ser Leu Gly Phe His Pro Ser Glu
249 35 40 45
251 Thr Pro Tyr Ile Lys Tyr Leu Glu Gly Ser Glu Thr Trp Lys Lys Val
252 50 55 60
254 Lys Leu Pro Thr Asp Gly Ile Ser Ala Ser Lys Ile Leu Gly Lys Ile
255 65 70 75 80
257 Met Ala Arg Val Arg Ile Ala Thr Ala Leu Ala Val Val Leu Ala Ala
258 85 90 95
260 Pro Cys Leu Ala Phe Asp Glu Val Thr Ala Ser Gly Val Phe Pro Glu
261 100 105 110
263 Glu His Lys His Thr Gly Glu Gly Arg His Leu Gln Thr Cys Thr Asn
264 115 120 125
266 Ser Asp Asp Ala Leu Asp Pro Thr Ala Pro Asn Arg Arg Asp Asn Val
267 130 135 140
269 Ala Phe Ala Ser Arg Arg Asp Ala Ala Arg Arg Glu Arg Asp Gly Thr
270 145 150 155 160
272 Gly Thr Val Cys Gln Ile Thr Asn Gly Glu Thr Asp Leu Ala Thr Met
273 165 170 175
275 Phe His Lys Ser Leu Pro His Asp Glu Leu Gly Gln Val Thr Ala Asp
276 180 185 190

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/840,762

DATE: 06/08/2001

TIME: 14:15:18

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\1840762.raw

```

278 Asp Phe Ala Ile Leu Glu Asp Cys Ile Leu Asn Gly Asp Phe Ser Ile
279      195      200      205
281 Cys Glu Asp Val Pro Ala Gly Asp Pro Ala Gly Arg Leu Val Asn Pro
282      210      215      220
284 Thr Ala Ala Phe Ala Ile Asp Ile Ser Gly Pro Ala Phe Ser Ala Thr
285 225      230      235      240
287 Thr Ile Pro Pro Val Pro Thr Leu Ser Ser Pro Glu Leu Ala Ala Gln
288      245      250      255
290 Leu Ala Glu Leu Tyr Trp Met Ala Leu Ala Arg Asp Val Pro Phe Met
291      260      265      270
293 Gln Tyr Gly Thr Asp Glu Ile Thr Thr Thr Ala Ala Ala Asn Leu Ala
294      275      280      285
296 Gly Met Gly Gly Phe Pro Asn Leu Asp Ala Val Ser Ile Gly Ser Asp
297      290      295      300
299 Gly Thr Val Asp Pro Phe Ser Gln Leu Phe Arg Ala Thr Phe Val Gly
300 305      310      315      320
302 Val Glu Thr Gly Pro Phe Val Ser Gln Leu Leu Val Asn Ser Phe Thr
303      325      330      335
305 Ile Asp Ala Ile Thr Val Glu Pro Lys Gln Glu Thr Phe Ala Pro Asp
306      340      345      350
308 Leu Asn Tyr Met Val Asp Phe Asp Glu Trp Leu Asn Ile Gln Asn Gly
309      355      360      365
311 Gly Pro Pro Ala Gly Pro Glu Glu Leu Asp Glu Glu Leu Arg Phe Ile
312      370      375      380
314 Arg Asn Ala Arg Asp Leu Ala Arg Val Ser Phe Val Asp Asn Ile Asn
315 385      390      395      400
317 Thr Glu Ala Tyr Arg Gly Ser Leu Ile Leu Leu Glu Leu Gly Ala Phe
318      405      410      415
320 Ser Arg Pro Gly Ile Asn Gly Pro Phe Ile Asp Ser Asp Arg Gln Ala
321      420      425      430
323 Gly Phe Val Asn Phe Gly Thr Ser His Tyr Phe Arg Leu Ile Gly Ala
324      435      440      445
326 Ala Glu Leu Ala Gln Arg Ala Ser Cys Tyr Gln Lys Trp Gln Val His
327      450      455      460
329 Arg Phe Ala Arg Pro Glu Ala Leu Gly Gly Thr Leu His Asn Thr Ile
330 465      470      475      480
332 Ala Gly Asp Leu Asp Ala Asp Phe Asp Ile Ser Leu Leu Glu Asn Asp
333      485      490      495
335 Glu Leu Leu Lys Arg Val Ala Glu Ile Asn Ala Ala Gln Asn Pro Asn
336      500      505      510
338 Asn Glu Val Thr Tyr Leu Leu Pro Gln Ala Ile Gln Val Gly Ser Pro
339      515      520      525
341 Thr His Pro Ser Tyr Pro Ser Gly His Ala Thr Gln Asn Gly Ala Phe
342      530      535      540
344 Ala Thr Val Leu Lys Ala Leu Ile Gly Leu Asp Arg Gly Gly Glu Cys
345 545      550      555      560
347 Phe Pro Asn Pro Val Phe Pro Ser Asp Asp Gly Leu Glu Leu Ile Asn
348      565      570      575
350 Phe Glu Gly Ala Cys Leu Thr Tyr Glu Gly Glu Ile Asn Lys Leu Ala

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/840,762

DATE: 06/08/2001

TIME: 14:15:19

Input Set : C:\PAOLA\09840762.txt

Output Set: N:\CRF3\06082001\I840762.raw